

주성분 분석과 로지스틱 회귀분석을 이용한 다국 통화 포트폴리오 전략

심경식¹ · 안재준² · 오경주³

¹²³연세대학교 정보산업공학과

접수 2011년 11월 21일, 수정 2011년 12월 24일, 게재확정 2012년 1월 17일

요약

본 논문에서는 외환시장에서 주성분 분석과 로지스틱 회귀분석을 이용한 다국 통화 포트폴리오 전략을 개발하는 것을 제안한다. 과거 환율시장의 분석에 대한 많은 연구가 진행되어 왔으나 상대적으로 외환시장에서의 거래 전략을 개발하는 연구는 거의 없었다. 본 연구는 크게 두 가지 목적을 가지고 있다. 첫 번째 목적은 주성분 분석을 적용시켜 포트폴리오를 구성하는 다양한 나라의 환율에 가중치 할당 방법을 제안하는 것이다. 두 번째 목적은 로지스틱 회귀분석을 이용하여 구성된 포트폴리오의 적절한 매수시점과 매도시점을 정하는 것이다. 이 논문의 실험결과는 제안한 투자전략의 유용성을 증명할 수 있을 것이며, 또한 이를 통해 시장참여자들에게 투자 결정에 있어 도움을 줄 수 있을 것이다.

주요용어: 다국 통화 포트폴리오, 로지스틱 회귀분석, 외환시장, 주성분 분석.

1. 서론

외환시장은 전 세계적으로 가장 활발히 거래되고 있는 시장 중 하나로써 많은 투자자들이 이 시장에 참여하고 있다. 세계 주요외환시장의 거래를 24시간 연계시키면서 모든 시장정보를 환율에 신속하게 그리고 지속적으로 반영하는 하나의 범세계적 시장으로서의 기능을 수행하면서 국제금융거래의 효율화를 촉진하고 있다.

이런 외환시장에 대한 연구는 현대 금융시장을 분석함에 있어 최근에 가장 각광받는 분야이며 (Beran과 Ocker, 1999; Fernando 등, 1999), 시장 참여자들에게 더 정확한 정보를 제공할 수 있도록 많은 연구자들에게서 다양한 환율예측 모형이 제안되고 있다 (Gencay, 1997). 특히 1990년대와 2000년 후반에 발생되어 많은 개발 도상국들을 휩쓸었던 금융위기와 미국 서브프라임 모기지를 시작으로 발생한 금융위기는 외환시장의 변동성을 크게 증가시킨 계기가 되었다. 이러한 사건들을 계기로 외국 환율 시장의 변동성 증가로 인해 금융시장에서 특히 환율 시장에 대한 관심은 점점 높아지고 있다.

Cushman (2007)은 캐나다-미국 환율에 대한 Frankel (1984)의 모델을 수정한 밸런스 포트폴리오 모델을 제안했다. 그는 포트폴리오 밸런스 모델의 변수들 사이에서 공적분 (cointegration)을 발견하였고, 통합된 벡터는 넓게는 이론적 기대와 일치한다는 사실을 밝혀냈다 (Frankel, 1984). Pai 등 (2010)은 러프셋 이론 (rough set theory)과 비순환 그래프 SVM (acyclic graph support vector machines)을 통

¹ (120-749) 서울특별시 서대문구 신촌동 134번지, 연세대학교 정보산업공학과, 석사과정.

² (120-749) 서울특별시 서대문구 신촌동 134번지, 연세대학교 정보산업공학과, 박사과정.

³ 교신저자: (120-749) 서울특별시 서대문구 신촌동 134번지, 연세대학교 정보산업공학과, 부교수.

E-mail: johanoh@yonsei.ac.kr

해 외국 환율을 분석하였다. Huang 등 (2011)은 환율의 불규칙한 수익 패턴을 강조하였고 분위회귀분석 방법 (quantile regression method)으로 환율의 변동성을 예측하였다. 그리고 Seemann 등 (2011)은 가장 활발한 외국 환율 시장에 대한 분산의 변동성 증가를 분석하였으며, Qui 등 (2011)은 구매력 평가 (purchasing power parity)를 이용하여 외국 환율 움직임을 예측하는 모델을 제시하였다. 그들은 역사적 편차 (historic deviations)를 사용하여 장기간의 환율 정보를 통해 자국 통화와 외국 환율 사이의 관계를 중점적으로 설명하였다. Bahmani-Oskooee와 Harvey (2011)은 미국과 말레이시아 사이의 환율의 변동성과 산업 간의 투자를 연구하였다. 그들은 환율의 변동성이 말레이시아의 투자 흐름에 어떠한 영향을 주는지 조사하였다.

그러나 이처럼 기존의 외환시장에 대한 연구들은 환율을 예측하거나 환율 시장의 매커니즘을 분석하는데 그쳤다. 즉, 환율 시장의 분석에 대한 연구는 기존에 많이 진행되고 있으나 외환시장에서의 거래 전략들을 개발하는 연구는 거의 없는 상황이다. 따라서 이 연구를 통해 주성분 분석과 로지스틱 회귀분석을 기반으로 한 통합 모델을 제시하고, 다국 통화 포트폴리오 전략을 제시하고자 한다. 이 연구는 크게 두 가지 목적을 가지고 있다. 첫 번째 목적은 주성분 분석을 적용시킨 새로운 포트폴리오 할당 방법을 제안하는 것이며, 두 번째 목적은 로지스틱 회귀분석을 통하여 적절한 매수시점과 매도시점을 결정하는 것이다. 다시 말하면, 다국 통화 움직임을 설명하는 성분을 추출하고, 매수 포지션에서 매도 포지션으로의 이동이나 매도 포지션에서 매수 포지션으로의 이동과 같은 포지션 변경 시에 돌파 시점, 임계값 (threshold point), 전환 시점을 결정하고자 하는 것이다.

본 논문은 다음과 같이 구성되어 있다. 2절에서는 우리가 제안하는 투자전략을 자세히 설명하였으며, 3절에서는 제안한 투자전략의 실증분석 결과를 분석하였으며, 마지막으로 결론에서는 본 연구의 기대효과 및 향후 연구에 대해 서술하였다.

2. 제안한 투자전략

본 절에서는 제안한 투자전략의 구조와 특징에 대해 설명한다. 제안한 투자전략의 1 단계에서는 다국 통화 움직임을 설명하는 성분을 추출하고 다국 통화 포트폴리오 가중치를 결정된 뒤 성분 점수를 도출한다. 2 단계에서는 주성분의 임계값을 결정한다. 그리고 마지막 단계에서는 로지스틱 회귀분석을 통한 최종 다국 통화 포트폴리오 거래전략을 개발 한다.

1 단계. 다국 통화 움직임을 설명하는 성분 추출과 가중치 결정 및 성분 점수 도출

다국 통화 포트폴리오 전략을 위해서는 주성분의 시계열이 시간에 대하여 가격 움직임을 포트폴리오를 나타내는 지표로 사용되어야 한다. 또한 이러한 사용은 다국 통화들로 구성된 주성분의 집합을 위한 각각의 고유벡터 e_i 는 다국 통화 포트폴리오에 영향을 주는 각각의 변수들에게 최적의 상태로 가중치가 할당된다. 각각의 고유벡터 선형결합으로 통하여 각각의 성분 점수를 얻을 수 있고 이는 $y_i = e_i'x$ 와 같다 (여기서 x 는 9개국 환율 데이터). 각각의 고유벡터는 단순히 각각의 선형결합 매개 변수를 나타낸다. 그러나 각각의 포트폴리오 가중치는 $e_i'e_j = 1$ ($i = j$)라는 조건에 의해 결정되어야 하며, 포트폴리오 할당은 가중치로 결정된다.

2 단계. 주성분의 임계값 (threshold points) 결정

각각의 환율에 대한 가격은 시간에 따라서 변하고 있다. 환율 거래에서 흥미로운 것은 이런 변화들이 주어진 기간에서 증가하고 감소하는 일련의 패턴을 가지고 있는 점이다. 식 (2.1)과 (2.2)를 통해 임계값 (threshold points)을 결정한다. 매수 포지션이나 매도 포지션을 결정하기 위해서는 적용될 로지스

틱 회귀분석을 통해 임계값을 뚫고 올라가거나 내려갔을 때 포지션을 취하게 된다.

$$g\left(\frac{p_t - p_{t-i}}{p_{t-i}}\right) = 1, \text{ 만일 } \frac{p_t - p_{t-i}}{p_{t-i}} > L, t \neq i \quad (2.1)$$

$$g\left(\frac{p_t - p_{t-i}}{p_{t-i}}\right) = -1, \text{ 만일 } \frac{p_t - p_{t-i}}{p_{t-i}} < -L, t \neq i \quad (2.2)$$

여기서 $g((p_t - p_{t-i})/p_{t-i})$ 는 매수 또는 매도 신호를 발생시키는 함수로써, 만일 $(p_t - p_{t-i})/p_{t-i} > L, t \neq i$ 이면 $g((p_t - p_{t-i})/p_{t-i})$ 값은 1로 매수 신호를, $(p_t - p_{t-i})/p_{t-i} < -L, t \neq i$ 이면 $g((p_t - p_{t-i})/p_{t-i})$ 값은 -1로 매도 신호를 발생시킨다. $-L < (p_t - p_{t-i})/p_{t-i} < L, t \neq i$ 일 때 $g((p_t - p_{t-i})/p_{t-i})$ 값이 0으로 매수 또는 매도 신호를 발생시키지 않는다. L 은 백분율로 나타내는 움직임의 크기, p_t 는 t 시점의 가격, p_{t-i} 는 $t-i$ 시점의 가격을 나타낸다.

3 단계. 로지스틱 회귀분석을 통한 투자 시스템 개발식 (2.3)에 나타나 있는 상대적 위치 (relative position)는 독립변수로써 2 단계에서 결정한 임계값 (threshold points)을 통하여 매수 포지션과 매도 포지션을 결정하는데 도움을 준다. 로지스틱 회귀분석을 통해 독립변수인 상대적 위치를 이용하여 매수 포지션과 매도 포지션을 결정한다. 포지션을 결정하는 종속변수는 임계값에 의하여 재설정 될 수 있다.

$$x_k = \frac{p_t - \pi_k}{\eta_k - \pi_k} \quad (2.3)$$

위 식에서 x_k 는 k 투자기간 동안 첫 번째 성분 점수의 상대적 위치를 나타내는 지표 (indicator), p_t 는 t 시점의 증가, π_k 는 k 투자기간 동안 최저 가격, η_k 는 k 투자기간 동안 최고 가격을 의미한다.

$$\log(odds) = 1, \text{ 만일 } g\left(\frac{p_t - p_{t-i}}{p_{t-i}}\right) = 1 \quad (2.4)$$

$$\log(odds) = 0, \text{ 만일 } g\left(\frac{p_t - p_{t-i}}{p_{t-i}}\right) = -1 \quad (2.5)$$

여기서 $\log(odds)$ 은 로짓 (logit)을 의미하고, 승산 (odds)은 집단 1에 속하는 확률에서 집단 0에 속하는 확률을 나눈 비율을 말한다. 분류기준값 (cut-off value)에 의해 0 (매도 포지션)과 1 (매수 포지션)의 이항형의 값을 가진다. 만약 분류기준값이 0.5 이하이면 매도 포지션으로 0값을 가지며, 0.5 초과이면 매수 포지션으로 1값을 가지게 된다.

3. 실증분석

본 연구에서 제안한 투자전략은 총 3 단계로 구성되어 있다 (그림 3.1). 이 연구에 사용된 데이터는 외환시장에서의 전체 거래량 중 거래량의 비중이 높은 9개국의 통화를 선택하였고, 2005년 1월 1일부터 2009년 12월 31일까지의 일 데이터를 사용하였으며, 이 데이터들은 한국은행 경제통계시스템 (<http://ecos.bok.or.kr>)에서 얻을 수 있다.

표 3.1에서 볼 수 있는 여러 나라의 통화 중 홍콩 달러 (HKD)의 경우 의도적으로 배제하였는데, 그 이유는 영국 통치의 국가에서 중국 정부가 1997년부터 통치한 이후 정치적·사회적 제제로 인해 환율의 움직임에 신뢰성이 떨어진다고 판단했기 때문이다. 표 3.1은 외환시장에서의 외국 환율의 일별 투자 비율이다.

미국 달러 대비 9개국 통화들은 주성분 분석을 통하여 작은 차원으로 줄일 수 있다. 경험적으로도 이론적으로도 성분의 적절한 수를 결정하는 것은 컷-오프 방식 (cut-off method)으로도 완벽할 수 없다. 그러나 각각의 고유값에 대한 비율은 예전부터 결정되고 있고, 그 비율은 시계열 자료의 움직임을 설명

할 수 있는 설명력을 가지고 있다. 이 논리에 따르면, 첫 번째 성분의 경우 분산의 66% 설명할 수 있고, 두 번째 성분의 경우 분산의 26%를 설명할 수 있다 (표 3.2).

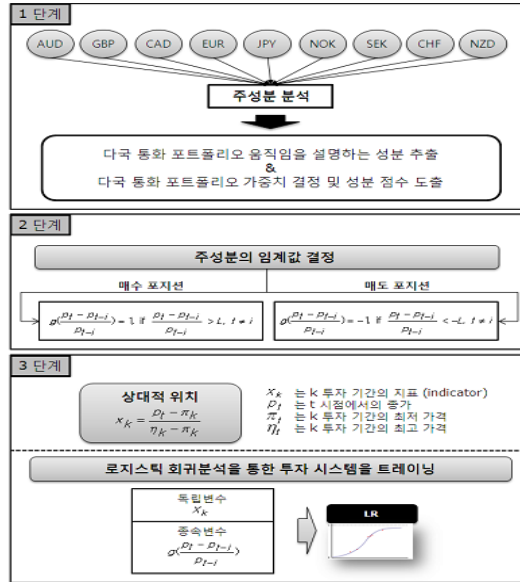


그림 3.1 본 연구의 제안된 투자전략

표 3.1 외환시장에서의 거래량 순위

순위	통화	% 일별 점유율
1	US dollar (USD)	43.15
2	Euro (EUR)	18.5
3	Yen (JPY)	8.25
4	Pound Sterling (GBP)	7.5
5	Swiss Franc (CHF)	3.4
6	Australian Dollar (AUD)	3.35
7	Canada Dollar (CAD)	2.1
8	Swedish Krona (SEK)	1.4
9	Hong Kong Dollar (HKD)	1.4
10	Norwegian Krone (NOK)	1.1
11	New Zealand Dollar (NZD)	0.95
12	Mexican Peso (MXN)	0.65
13	Singapore Dollar (SGD)	0.6
14	Won (KRW)	0.55
15	Others	7.1
총 합계		100

Bank for International Settlements, 2007년 12월)

표 3.2 고유값과 설명된 비율의 분산

구분	주성분1	주성분2	주성분3	주성분4	주성분5	주성분6	주성분7	주성분8	주성분9
고유값	5.87	2.36	0.35	0.25	0.06	0.04	0.03	0.02	0.01
설명된 분산	66%	26%	4%	3%	1%	0%	0%	0%	0%

첫 번째 주성분의 설명된 분산 값이 66%라는 충분히 높은 설명력을 가진다는 것은 본 연구에서 첫 번

째 성분 점수로 부터 다국 통화 포트폴리오의 매매 시점을 결정하는데 무리가 없다는 것을 보여준다고 할 수 있다. 표 3.3는 첫 번째 주성분의 고유벡터와 포트폴리오 가중치를 보여주고 있다. 각각의 포트폴리오 구성요소는 주성분 분석의 제약 조건인 $e_i^T e_j = 1 (i = j)$ 에 의해 가중치가 결정되고, 이는 전통적인 포트폴리오 최적기법과 비교하여 전혀 다르다. 이 고유벡터는 변수들 사이에서 독립적으로 예상된 포지션 뿐 만아니라 각각의 포트폴리오 성분이 첫 번째 주성분의 움직임에 얼마나 영향을 주는지에 대해 설명하는 포트폴리오 가중치로도 나타낼 수 있다.

표 3.3을 보면 알 수 있듯이 첫 번째 주성분의 고유벡터는 모두 양의 값을 가진다. 이 9개국 통화는 첫 번째 주성분에 양의 영향을 주기 때문에 매수의 포지션을 취함을 의미한다. 하지만 만약 어느 통화의 고유벡터의 값이 음의 값을 가진다면 매도 포지션을 취해야 한다.

표 3.3 고유벡터와 포트폴리오 가중치

통화	AUD	GBP	CAD	EUR	JPY	NOK	SEK	CHF	NZD
성분 1	0.40	0.26	0.37	0.36	0.01	0.40	0.39	0.28	0.33
PC1의 가중치	15.9%	6.9%	13.4%	13.1%	0%	16.3%	15.4%	7.8%	11.2%

첫 번째 주성분의 가중치로 각국 환율 데이터에 적용시켜 다국 통화 포트폴리오를 구성한다. 그림 3.2는 실험기간 동안 첫 번째 성분 점수의 흐름을 보여준다. 주성분의 임계값을 결정하기 위하여 최적 투자기간과 L 의 값을 선정하였다. 표 3.4는 투자기간에 따른 연도별 수익률을 나타내며 투자기간 중 55일이 가장 높은 수익률을 보였기 때문에 이 연구의 최적 투자기간으로 결정하였다. L 의 값은 보통 경험적으로 2-4%로 결정되는데 2%의 경우 거래 빈도가 높고 수익률이 떨어지며, 반대로 4%의 경우는 거래의 빈도가 낮고 거래의 시점을 잘 파악하지 못하였다. 그 결과 여기서 L 의 값은 3%로 정하였다.

표 3.4 투자기간에 따른 연도별 수익률

투자기간	수익률 (%)		
	35일	55일	85일
2005	-4.73%	4.16%	-2.31%
2006	-0.72%	9.29%	5.94%
2007	5.12%	5.97%	-0.01%
2008	-5.04%	3.85%	0.05%
2009	7.07%	15.85%	0.98%

가장 높은 수익률을 보인 최적 투자기간 55일을 이용하여 $(p_t - p_{t-i})/p_{t-i}$ 의 변화율을 계산한다. 즉, $(p_t - p_{t-i})/p_{t-i}$ 은 55일 투자기간 동안의 변화율이 된다 ($t = 56, i = 55$). 55일 투자기간 동안의 상승률인 L 의 값이 3% 이상이면 $g((p_t - p_{t-i})/p_{t-i}) = 1$ 로 매수 신호, 반대로 55일 투자기간 동안의 하락률인 $-L$ 의 값이 -3% 이하이면 $g((p_t - p_{t-i})/p_{t-i}) = -1$ 로 매도 신호를 발생시킨다.

로지스틱 회귀분석을 사용하기 위한 독립변수는 상대적 위치 $x_k = (p_t - \pi_k)/(\eta_k - \pi_k)$ 로 나타내고 (여기서 p_t 는 55일이 되는 시점의 종가, π_k 는 55일 투자기간 동안 최저 가격, η_k 는 55일 투자기간 동안 최고 가격을 의미), 종속변수는 $g((p_t - p_{t-i})/p_{t-i})$ 가 된다. 로지스틱 회귀분석을 통하여 매수 포지션과 매도 포지션을 예측하는데 여기서는 분류기준값 (cut-off value)에 의해 0과 1의 이항형의 값을 가진다. 만약 분류기준값이 0.5 초과이면 $\log(odds) = 1$, 0.5 이하이면 $\log(odds) = 0$ 이 된다. 2 단계와 3 단계를 종합하여 종속변수 $g((p_t - p_{t-i})/p_{t-i}) = 1$ 로 매수 신호가 발생될 때 예측값이 $\log(odds) = 1$ 이면 매수 포지션, 반대로 종속변수 $g((p_t - p_{t-i})/p_{t-i}) = -1$ 로 매도 신호가 발생될 때 예측값이 $\log(odds) = 0$ 이면 매도 포지션을 취하게 된다.

표 3.5는 본 연구에서 제안한 다국 통화 포트폴리오 거래 전략을 실험기간 동안 적용시킨 결과를 요

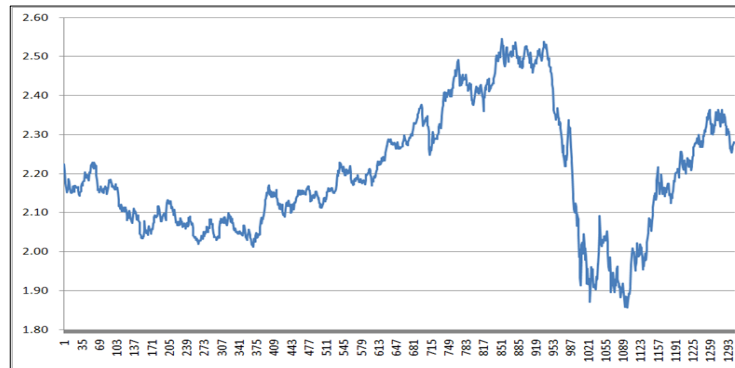


그림 3.2 9개국 통화들의 첫 번째 성분 점수

약하여 보여주고 있다. 전체 거래횟수 중 수익이 난 횟수의 비율은 74.5%로 매우 높은 수치의 winning ratio 값을 보여주고 있다. 이는 제안한 투자전략의 안정성을 보여주는 결과라고 판단할 수 있으며, 거래 당 최대 손실 또한 -2.0%인 사실을 감안한다면 위험 회피자 (risk averter)의 성향을 가지는 대부분의 투자자들에게 매력적인 투자전략이 될 수 있음을 알 수 있다.

표 3.5 실험기간 동안 제안한 투자전략의 거래 결과 요약

총 트레이딩 횟수	수익거래비율	거래당 최대수익	거래당 최대손실	평균 보유 기간
94	74.5%	3.8%	-2.0%	17.5 일

표 3.6은 실험기간 동안 연도별 수익률을 보여주고 있는데, 실험기간인 5년 동안 누적 수익률은 39.12%이며 연평균 수익률은 7.82%를 나타내고 있다. 특히 연평균 7.82%라는 수익률 수치는 시장 이자율이 약 4% 내외라는 것을 고려하였을 때 제안한 투자전략이 약 두 배 정도의 수익을 가져다줌으로서 실제로도 시장에서 매우 유용한 투자 전략이 될 수 있음을 보여준다. 따라서 이러한 실험결과를 통해, 본 연구에서 제안하는 다국 통화 포트폴리오 거래 전략은 비교적 안정적이며 높은 수익률을 보장하므로 많은 투자자들에게 (특히, 위험회피자의 특징을 가지는 투자자) 매우 유용한 대안을 제시할 수 있을 것이다. 그림 3.3은 제안한 투자전략을 통한 누적 거래 수익률을 그래프로 보여주고 있다.

표 3.6 실험기간 동안 제안한 투자전략의 연도별 수익률과 누적 및 평균 수익률

	2005	2006	2007	2008	2009	누적	평균
수익률 (%)	4.16%	9.29%	5.97%	3.85%	15.85%	39.12%	7.82%

4. 결론

외환시장은 전 세계에서 가장 큰 금융 시장 중 하나로서, 이 시장에서의 거래 전략 개발은 금융관련 전문가들에게 매우 주목받는 분야로 인식되고 있다. 따라서 본 연구는 주성분 분석과 로지스틱 회귀분석을 사용하여 외환시장에서의 투자 전략을 개발하고자 하였다. 제 3절에서 보여준 실험 결과로부터 본 연구에서 제안한 투자전략은 변동성이 심한 외환시장에서 투자자들에게 비교적 안정적이면서 높은 수익률을 낼 수 있도록 도움을 주는 거래 전략을 제공하기에 충분하다고 판단할 수 있었다. 본 연구에는

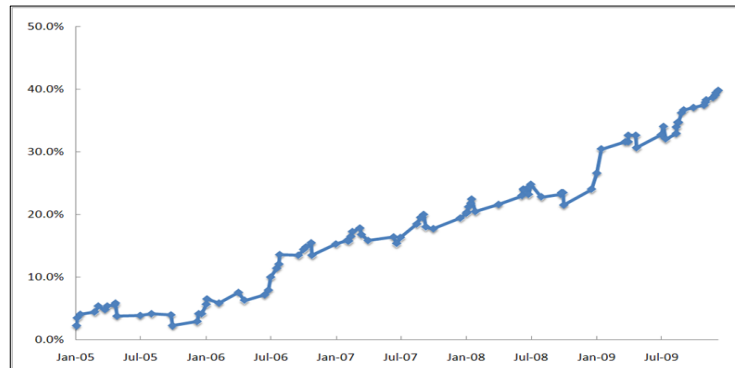


그림 3.3 제한한 투자전략의 누적 거래 수익률

9개국 환율의 일 데이터를 사용함으로써 제한한 투자전략을 통해 다양한 환율 데이터에 제약 받지 않을 수 있는 유용한 투자전략을 제시하였다. 이를 통해 본 연구결과가 많은 투자자들에게 투자 결정을 함에 있어서 중요한 정보를 제공할 수 있을 것으로 기대해 본다.

본 연구에서는 전통적인 통계기법만을 이용하여 투자전략을 구축하였다. 하지만 금융시장에서 획득되는 데이터들은 대부분 비선형적인 성격을 가진다는 사실을 고려하였을 때 추후 연구에서는 최근 금융 시장 분석 기법으로 적합하다고 알려진 인공지능 알고리즘을 추가로 고려해야 된다고 판단된다. 대표적인 인공지능 기법으로 알려진 인공신경망 (Artificial Neural Network; ANN)과 의사결정나무 (Decision Tree; DT)와 같은 다양한 분류 알고리즘을 사용한다면 더 좋은 결과가 나타날 수 있기 때문에 이러한 다양한 분석을 통해 추후 연구에서는 외환시장에서의 수익률을 극대화할 수 있는 더 나은 투자 전략 시스템이 나올 수 있기를 기대한다.

참고문헌

- 강명욱, 김부용, 홍주희 (2010). 로지스틱모형에서 그래픽을 이용한 회귀와 모형평가. <한국데이터정보과학회지>, **21**, 21-32
- 권세혁 (2010). 시뮬레이션 실험조건 이상 진단 연구. <한국데이터정보과학회지>, **21**, 853-861
- 권오진, 김태윤, 송규문 (2010). 붓스트랩 기법을 이용한 환율의 장단기 신뢰구간 예측. <한국데이터정보과학회지>, **21**, 493-502
- 김윤대, 전치혁, 이혜선 (2011). 별점 부분최소자승법을 이용한 분류방법. <한국데이터정보과학회지>, **22**, 931-940
- 김태윤, 권오진 (2011). 경제위기시 환율신뢰구간 예측 알고리즘 개발. <한국데이터정보과학회지>, **22**, 895-902
- 변현우, 송치우, 한성권, 이태규, 오경주 (2009). 변동성 지수기반 유전자 알고리즘을 활용한 계층구조 포트폴리오 최적화에 관한 연구. <한국데이터정보과학회지>, **20**, 467-478.
- Bahmani-Oskooee, M. and Harvey, H. (2011). Exchange-rate volatility and industry trade between the U.S. and Malaysia. *Research in International Business and Finance*, **25**, 127-155.
- Beran, J. and Ocker, D. (1999). SEMIFAR forecasts, with applications to foreign exchange rates. *Journal of Statistical Planning and Inference*, **80**, 137-153.
- Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and regression trees*, Wadsworth Intl Group, Belmont.
- Cherkassky, V. and Mulier, F. (1998). *Learning from data : Concepts, theory and methods*. Wiley, New York.
- Cushman, D. O. (2007). A portfolio balance approach to the Canadian-U.S. exchange rate. *Review of Financial Economics*, **16**, 305-320.

- Fernando, F. R., Simon, S. R. and Julian, A. F. (1999). Exchange rate forecasts with simultaneous nearest-neighbor methods: Evidence from the EMS. *International Journal of Forecasting*, **15**, 383-392.
- Frankel, J. A. (1984). Tests of monetary and portfolio balance models of exchange rate determination. In *Exchange rate theory and practice*, edited by J. Bilson, R. Marston, University of Chicago Press, Chicago.
- Gencay, R. (1999). Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules. *Journal of International Economics*, **47**, 91-107.
- Huang, A. Y., Peng, S. P., Li, F. and Ke, C. J. (2011). Volatility forecasting of exchange rate by quantile regression. *International Review of Economics and Finance*, **20**, 591-606.
- McCullagh, P. and Nelder, J. (1989). *Generalized linear models*, second edition, Chapman and Hall/CRC, Boca Raton.
- Pai, P. F., Chen, S. Y., Huang, C. W. and Chang, Y. H. (2010). Analyzing foreign exchange rates by rough set theory and directed acyclic graph support vector machines. *Expert Systems with Applications*, **37**, 5993-5998.
- Quinlan, J. R. (1986). Induction of decision tree. *Machine Learning*, **1**, 81-106.
- Seemann, L., McCauley, J. L. and Gunaratne G. H. (2011). Intraday volatility and scaling in high frequency foreign exchange markets. *International Review of Financial Analysis*, **20**, 121-126.
- Song, G. M., Park, B. C. and Kang, H. K. (2007). A CUSUM algorithm for early detection of structural changes in won/dollar exchange market. *Journal of the Korean Data & Information Science Society*, **18**, 345-356.
- Tarumi, T., Tanaka, Y. and Shin, J. K. (1991). Sensitivity analysis in principal component regression. *Journal of the Korean Data & Information Science Society*, **2**, 1-9.
- White. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, **2**, 359-366.
- Zhang, G. Q. and Michael, Y. H. (1998). Neural network forecasting of the British pound/US dollar exchange rate. *Omega*, **26**, 495-506.

Multi-currencies portfolio strategy using principal component analysis and logistic regression

Kyung Sik Shim¹ · Jae Joon Ahn² · Kyong Joo Oh³

¹²³Department of Information and Industrial Engineering, Yonsei University

Received 21 November 2011, revised 24 December 2011, accepted 17 January 2012

Abstract

This paper proposes to develop multi-currencies portfolio strategy using principal component analysis (PCA) and logistic regression (LR) in foreign exchange market. While there is a great deal of literature about the analysis of exchange market, there is relatively little work on developing trading strategies in foreign exchange markets. There are two objectives in this paper. The first objective is to suggest portfolio allocation method by applying PCA. The other objective is to determine market timing which is the strategy of making buy or sell decision using LR. The results of this study show that proposed model is useful trading strategy in foreign exchange market and can be desirable solution which gives lots of investors an important investment information.

Keywords: Foreign exchange market, logistic regression, multi-currencies portfolio, principal component analysis.

¹ Graduate student, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea.

² Ph.D. candidate, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea.

³ Corresponding author: Associate professor, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea. E-mail: johanoh@yonsei.ac.kr